

# RNA polymerase II senses obstruction in the DNA minor groove via a conserved sensor motif

Liang Xu<sup>a,1</sup>, Wei Wang<sup>a,1</sup>, Deanna Gotte<sup>b</sup>, Fei Yang<sup>c</sup>, Alissa A. Hare<sup>c</sup>, Timothy R. Welch<sup>c</sup>, Benjamin C. Li<sup>c</sup>, Ji Hyun Shin<sup>a</sup>, Jenny Chong<sup>a</sup>, Jeffrey N. Strathern<sup>b,2</sup>, Peter B. Dervan<sup>c,2</sup>, and Dong Wang<sup>a,2</sup>

<sup>a</sup>Department of Cellular and Molecular Medicine, School of Medicine, Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego, La Jolla, CA 92093; <sup>b</sup>Gene Regulation and Chromosome Biology Laboratory, Frederick National Laboratory for Cancer Research, National Cancer Institute, Frederick, MD 21702; and <sup>c</sup>Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125

Contributed by Peter B. Dervan, September 12, 2016 (sent for review June 15, 2016; reviewed by Roger D. Kornberg and Jesper Svestrup)

**RNA polymerase II (pol II) encounters numerous barriers during transcription elongation, including DNA strand breaks, DNA lesions, and nucleosomes. Pyrrole-imidazole (Py-Im) polyamides bind to the minor groove of DNA with programmable sequence specificity and high affinity. Previous studies suggest that Py-Im polyamides can prevent transcription factor binding, as well as interfere with pol II transcription elongation. However, the mechanism of pol II inhibition by Py-Im polyamides is unclear. Here we investigate the mechanism of how these minor-groove binders affect pol II transcription elongation. In the presence of site-specifically bound Py-Im polyamides, we find that the pol II elongation complex becomes arrested immediately upstream of the targeted DNA sequence, and is not rescued by transcription factor IIS, which is in contrast to pol II blockage by a nucleosome barrier. Further analysis reveals that two conserved pol II residues in the Switch 1 region contribute to pol II stalling. Our study suggests this motif in pol II can sense the structural changes of the DNA minor groove and can be considered a “minor groove sensor.” Prolonged interference of transcription elongation by sequence-specific minor groove binders may present opportunities to target transcription addiction for cancer therapy.**

Py-Im polyamide | transcription inhibition | minor groove | DNA

In eukaryotes, precursor mRNA synthesis is catalyzed by the RNA polymerase II holoenzyme (pol II), which frequently pauses during transcription elongation (1–3). In addition to regulatory factors that control pol II processivity, various obstacles encountered by pol II can also lead to stalling, and even backtracking, on the DNA template (4). Factors that affect pol II transcription elongation dynamics include intrinsic DNA sequences or structures (5, 6), endogenous epigenetic DNA modifications (7), embedded ribonucleotides (8), DNA lesions (9, 10), small-molecule DNA-binders (11, 12), DNA-binding proteins including nucleosomes (13), and even pol II itself (14, 15). Transient transcriptional pausing or short-lived transcriptional blockage can be rescued by the recruitment of transcription factor IIS (TFIIS) (16, 17), a transcription factor that facilitates the cleavage of backtracked transcript. In contrast, prolonged transcriptional arrest by some bulky DNA lesions triggers either ubiquitination and degradation of pol II, or transcription-coupled nucleotide excision repair, a special DNA repair pathway that preferentially repairs DNA lesions in the transcribed strand (17, 18). Structural, genetic, and biochemical studies have greatly advanced our understanding of how pol II copes with different kinds of covalent DNA lesions caused by oxidation (19, 20), alkylation (21–24), and photocyclo-addition (10). Less well understood are the interactions of the pol II machinery when confronted with a steric blockage by small molecules bound noncovalently in the minor groove of DNA (25–27).

Pyrrole-imidazole (Py-Im) polyamides are a class of small molecules that can be programmed to selectively target specific DNA sequences in the minor groove with high binding affinity (28, 29). This sequence-specificity arises from unique pairs of aromatic amino acids, which distinguish the edges of the four Watson–Crick

base pairs (30, 31). Eight-ring hairpin oligomers linked by a central aliphatic  $\gamma$ -aminobutyric acid unit have binding affinity comparable to TFs (32). Early studies suggested that cell-permeable Py-Im polyamides modulate gene expression by targeting TF binding (25, 33, 34). Importantly, inhibition of pol II occupancy at gene start sites leading to Rpb1 degradation has been observed (35). The question arises as to how polyamide occupancy in the DNA minor groove prevents pol II elongation. Do Py-Im polyamides pose a direct impediment to pol II elongation and for how long? In this study, we investigate the mechanism of how pol II transcription elongation is affected by Py-Im polyamide binding at a discrete site downstream from the pol II active site. From molecular modeling, a conserved pol II motif appears to function as a minor groove sensor and could be responsible for pol II elongation inhibition. This finding was verified by site-specific mutations of this region in pol II. These results suggest direct interaction between residues of the conserved motif and the bound Py-Im polyamides. In short, we have identified a sensor function of a conserved motif in pol II.

## Results

### Py-Im Polyamides Are Roadblocks for Pol II Transcription Elongation.

To investigate the impact of a site-specific DNA-binding small molecule on pol II transcription elongation, we assembled a purified pol II elongation complex with a DNA scaffold (78 bp in

## Significance

**Transcription addiction is a hallmark of cancer and a potential therapeutic target. RNA polymerase II (pol II) is responsible for synthesizing precursor mRNA in all eukaryotic cells and can be blocked by obstacles, such as DNA lesions and nucleosomes on the DNA template. In this study, we demonstrate that sequence-specific minor groove binding pyrrole-imidazole polyamides can sterically block an elongating polymerase at the targeted binding site. We find this blockage is persistent and cannot be rescued by transcription factor IIS. We further show pyrrole-imidazole polyamides are detected in the minor groove via two conserved residues in the Switch 1 region of pol II. Collectively, these results provide mechanistic insights on how a noncovalent minor groove binder can obstruct pol II elongation.**

Author contributions: L.X., W.W., J.N.S., P.B.D., and D.W. designed research; L.X., W.W., D.G., and F.Y. performed research; A.A.H., T.R.W., B.C.L., J.H.S., and J.C. contributed new reagents/analytic tools; L.X., W.W., F.Y., J.N.S., P.B.D., and D.W. analyzed data; and L.X., W.W., F.Y., J.N.S., P.B.D., and D.W. wrote the paper.

Reviewers: R.D.K., Stanford University School of Medicine; and J.S., The Francis Crick Institute.

Conflict of interest statement: P.B.D. and F.Y. own shares in Gene Sciences, Inc., a biotechnology company that is developing therapeutics for prostate cancer.

Freely available online through the PNAS open access option.

<sup>1</sup>L.X. and W.W. contributed equally to this work.

<sup>2</sup>To whom correspondence may be addressed. Email: strathej@mail.nih.gov, dervan@caltech.edu, or dongwang@ucsd.edu.

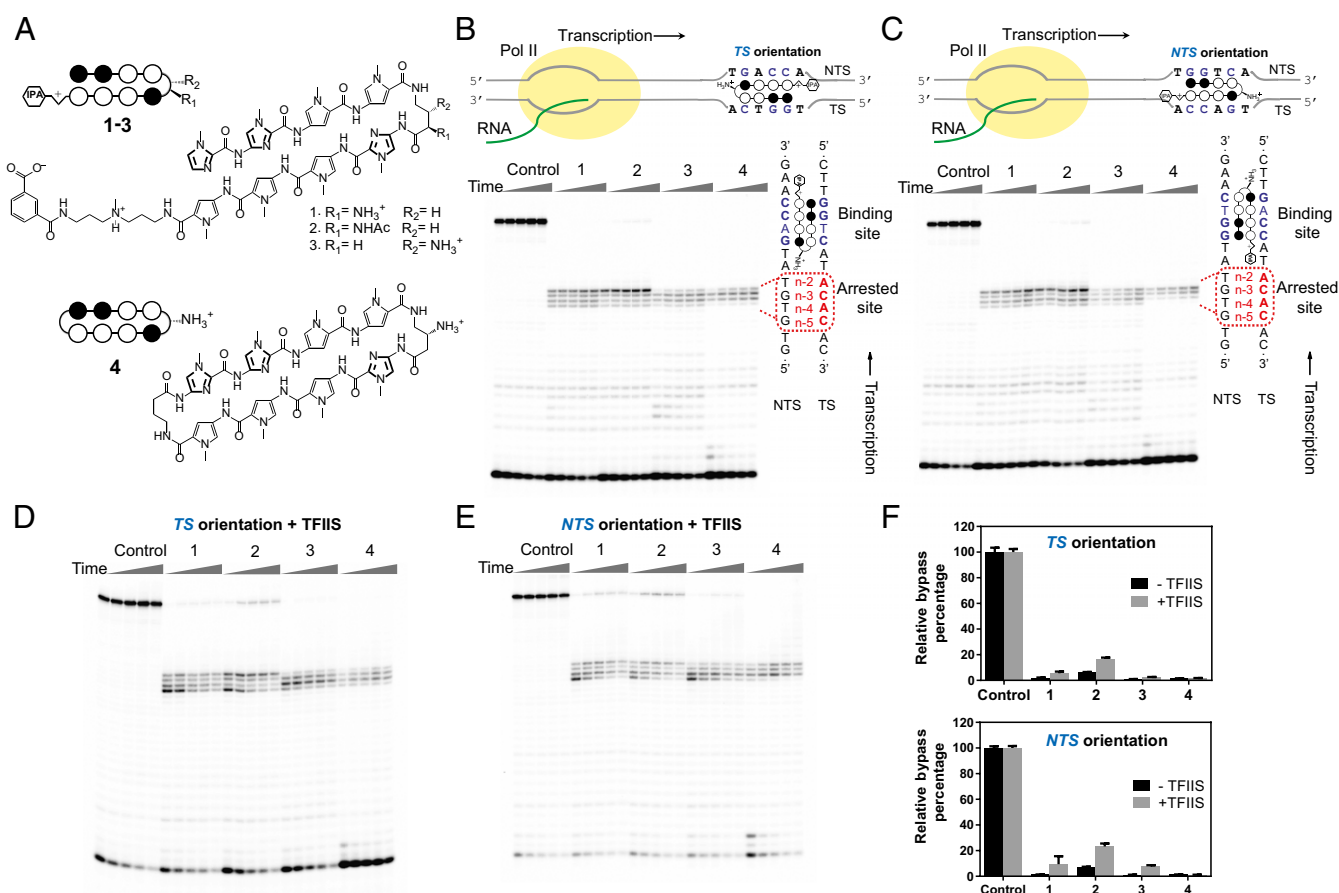
This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1612745113/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1612745113/-DCSupplemental).

length) containing a full transcription bubble upstream and a specific 6-bp binding site downstream (59 bp from the scaffold end). The 6-bp site is bound specifically by Py-Im polyamides 1–4 that code for 5'-WGGWCW-3' according to the pairing rules (Fig. 1). To address the potential differences of hairpin Py-Im polyamide binding orientation relative to transcriptional directionality (upper black arrow in Fig. 1B), we also designed a second scaffold with the downstream polyamide binding site in an opposite orientation (upper black arrow in Fig. 1C). We termed the orientation in which the pol II encounters the  $\gamma$ -turn moiety of the hairpin Py-Im polyamide first (Fig. 1B) the “template strand (TS) binding orientation,” and the reverse orientation where pol II would encounter the “C-terminal linker” first the “nontemplate strand (NTS) binding orientation” (Fig. 1C).

Substitutions on the  $\gamma$ -turn or the C terminus of Py-Im polyamides influence DNA affinity as well as cellular uptake (36). To evaluate the effects of these substitutions on pol II transcription elongation, we compared Py-Im polyamides 1–4 (Fig. 1A). Hairpins 1, 2, and 3 have identical C terminus and vary on the  $\gamma$ -turn, where 1 has a chiral (R)- $\alpha$ -amino group, 2 has an acetylated (R)- $\alpha$ -amino group, and 3 has a (R)- $\beta$ -amino substitution. Compound 4 is a cycle flanked by the same  $\beta$ -amino  $\gamma$ -turn as 3 and an unsubstituted  $\gamma$ -turn.

All four Py-Im polyamides (1–4) block RNA pol II transcription elongation at specific positions upstream of their binding

sites in both the TS and NTS orientations (Fig. 1B and C), whereas RNA pol II can transcribe through the DNA template to obtain full-length RNA transcripts in the absence of Py-Im polyamide treatment (see control lanes in Fig. 1B and C). Importantly, pol II cannot bypass the polyamide blockage even after prolonged incubation (2–20 h) (Fig. 1B and C and Fig. S1), implying that binding of Py-Im polyamides can induce prolonged arrest of pol II at the transcription elongation phase in a sequence-specific manner. To exclude the possibilities that pol II could fall off from DNA, or RNA could be released during the long time incubation in the presence of polyamides, we performed a native gel analysis to test stability of the polyamide-arrested pol II complex and we observed that the complex is very stable and no obvious RNA transcripts are released (Fig. S1C). To determine the stability of the polyamide-arrested pol II complex (without RNA transcript release) and whether transcription can recover upon the removal of polyamide, we designed competition experiments to remove the polyamide from the elongation complex by addition of competitor duplex DNA with the same binding site, and found that the elongation can be slowly recovered (18 h) and reach the full transcripts in the presence of high concentration of competitor duplex DNA (Fig. S1D). We found that the arrested region generally has two to four pausing bands that are located 2–5 bp upstream of the 5'-WGGWCW-3' binding site (see the dashed box on the right of the gel image in Fig. 1B and C, named



**Fig. 1.** Py-Im polyamides are strong roadblocks for RNA pol II transcription elongation. (A) Four Py-Im polyamides share the same 6-bp DNA target sequence (5'-WGGWCW-3'). The black solid circle refers to *N*-methylimidazole (Im), and the hollow circle refers to *N*-methylpyrrole (Py). (B) RNA pol II transcription blockage by Py-Im polyamides with a TS binding orientation and (C) with a NTS binding orientation. (D and E) PAGE gel analysis of TFIIS effect on Py-Im polyamide induced pol II arrest. The Py-Im polyamide binds DNA in a TS orientation (D) or NTS orientation (E). In these gel images, control lanes are RNA pol II transcription in the absence of Py-Im polyamides. Time points for transcription assays were 10 min, 20 min, 30 min, 1 h, and 2 h. The concentration of Py-Im polyamides was 0.5  $\mu\text{M}$ . (F) Quantitative analysis of TFIIS effect on rescuing Py-Im polyamide-induced pol II arrest.

the n-2, n-3, n-4, and n-5 bands, from top to bottom). These results reveal that pol II is able to sense the barrier and becomes arrested several base pairs upstream of the bound Py-Im polyamide.

Quantitative analysis of the pol II pausing bands reveals specific patterns of polyamide-induced pol II arrest. The pattern of pausing bands depends on the chemical nature of the polyamides, the concentration of the compounds, the binding orientations, and the incubation time (Fig. 1 and Figs. S1–S5). In the TS binding orientation, where the  $\gamma$ -turn faces toward oncoming pol II, acetylation of the  $\alpha$ -amine group in hairpin 2 causes more dominant pausing at the n-2 site than hairpin 1 (Fig. 1 and Fig. S1). Interestingly, moving the amine group into the  $\beta$ -position in hairpin 3 causes the major pausing sites to shift to n-3 and n-4 sites (Fig. 1 and Fig. S1). Finally, cycle 4 shares a similar pattern with hairpin 3 as might be expected, with a minor increase of the n-2 site with prolonged incubation (Fig. S1). In the NTS binding orientations where the C terminus faces toward the oncoming pol II, hairpins 1, 2, and 3 show similar stall patterns, which is consistent with their identical C-terminus moieties. In contrast, cycle 4 only has two major pausing sites (n-2 and n-3) and one minor pausing site (n-4), consistent with a shorter  $\gamma$ -turn facing pol II rather than a longer C-terminus linker. Despite the subtle pattern differences, both binding orientations had a common pausing region (n-2 to n-5) relative to the Py-Im polyamide binding site.

To further characterize the strength of Py-Im polyamides as inhibitors of pol II elongation, we performed concentration-dependent pol II inhibition analysis for the four polyamides 1–4 in both binding orientations (Figs. S2–S5). The full-length pol II transcripts are diminished and the arrested bands accumulate as polyamide concentration increases (Figs. S2–S5). To quantitatively evaluate the concentration-dependent inhibition data, we systematically determined the inhibition parameter  $K_i$  (the concentration of ligand that can inhibit 50% of pol II elongation) (see *SI Materials and Methods* for details) for each Py-Im polyamide and binding orientation (Table 1 and Figs. S2–S5). For each polyamide, the NTS binding orientation generally leads to a stronger inhibition effect than the TS binding orientation, suggesting an asymmetric inhibition of pol II transcription elongation. The  $K_i$  concentrations did not correlate with the binding affinity of these polyamides observed from duplex melting experiments (Table S1). A comparison of these four ligands reveals that 2 and 4 have the strongest inhibition. Comparison between hairpins 1, 2, and 3 clearly shows the influence of substitutions on the  $\gamma$ -turn, which faces the approaching enzyme in the TS binding mode. Acetylation of the  $\gamma$ -turn  $\alpha$ -amine group in 2, which decreases DNA binding affinity, increases the inhibition ability by ~threefold, whereas moving the  $\alpha$ -amine group to the  $\beta$ -position (3) decreases the inhibition ability by ~threefold. Interestingly, cyclization of 4 greatly increases the inhibition ability by ~16-fold over 3. Although modifications in the  $\gamma$ -turn and the C terminus can modulate the inhibition abilities of these molecules, all four molecules cause strong inhibition of pol II elongation with  $K_i$  values at nanomolar concentrations.

**TFIIS Is Insufficient to Fully Rescue Py-Im Polyamide-Arrested Pol II.** Transcription elongation factor TFIIS can rescue arrested pol II and enhance pol II bypass of a variety of translocation barriers, such as nucleosomes, some DNA lesions, small-molecule DNA binders, and intrinsic pausing DNA sequences (12, 13, 16, 37).

**Table 1. Concentration of Py-Im polyamides for 50% inhibition of RNA pol II bypass ( $K_i$ , nM)**

Orientation	1	2	3	4
TS	50 $\pm$ 10	~17 $\pm$ 2	160 $\pm$ 30	~10 $\pm$ 2
NTS	45 $\pm$ 12	<5	63 $\pm$ 17	~8 $\pm$ 2

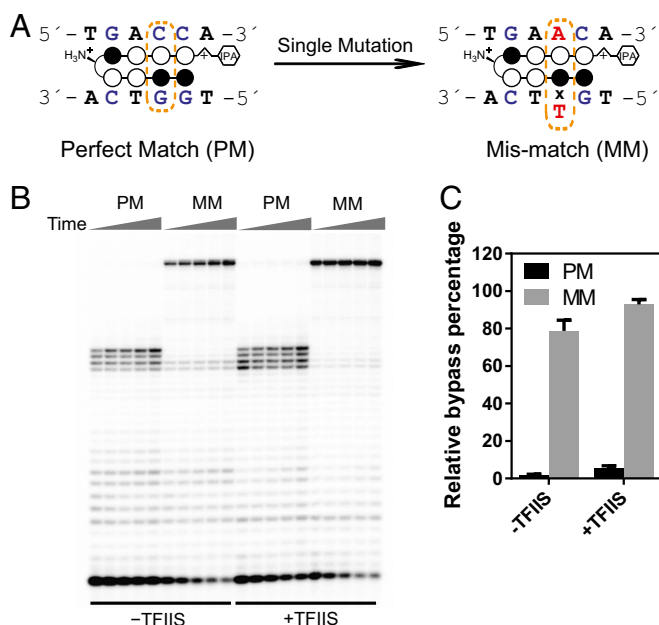
To investigate whether TFIIS can stimulate pol II bypass of polyamide-induced pol II arrest, we performed transcription elongation experiments in the absence or presence of TFIIS. As shown in Fig. 1 D and E, the majority of pol II remains arrested by Py-Im polyamides even in the presence of TFIIS, regardless of the binding orientation (TS or NTS), with only a small portion of transcriptional bypass observed for hairpins 1 and 2. These results indicate that TFIIS fails to effectively rescue polyamide-induced pol II arrest. The presence of TFIIS altered the patterns of the pausing bands, as the major arrested bands are shifted toward the upstream positions (Fig. 1 D and E, n-4 or n-5 positions). This finding suggests that pol II may adopt pretranslocation or backtracked states during Py-Im polyamide-induced arrest. TFIIS facilitates cleavage of short RNA oligos from the 3'-end of RNA transcripts, leading to the shortened transcripts and shifted pausing patterns.

Quantitative comparison of bypass percentage in the absence (Fig. 1 B and C) and presence of TFIIS (Fig. 1 D and E) reveals distinct responses to TFIIS among the four polyamides (Fig. 1F). For example, the lowest-affinity hairpin 2 is more sensitive to TFIIS treatment with ~12% transcript bypass of the blockage and ~88% arrest in the TS binding orientation. In contrast, the higher-affinity cycle 4 is resistant to TFIIS treatment. We observed no transcriptional bypass in the presence of TFIIS even after 2-h incubation. Similar TFIIS sensitivity profiles were also observed for NTS binding orientation, with cycle 4 most resistant to TFIIS treatment (Fig. 1F). These results demonstrate that TFIIS is insufficient to rescue Py-Im polyamide-induced pol II transcriptional arrest, indicating that Py-Im polyamide-induced transcriptional blockage is strong and stable. This finding is in contrast with other translocation barriers, such as nucleosomes or other small-molecule binders (Distamycin and DAPI) (12, 13).

**Sequence-Specific RNA Pol II Transcriptional Blockage by Py-Im Polyamides.** Our results support the DNA target specificity of Py-Im polyamides (32), as evidenced by these polyamides only causing pol II stalling specifically 2–5 bp upstream of the match binding site but not in other sites in the template DNA. To further directly verify that pol II transcription stalling by polyamides is indeed a result of its discrete binding to the 5'-WGGWCW-3' sequence, we compared the transcriptional inhibitory effect of hairpin 1 upon its binding to DNA with a perfectly matched (PM) sequence (5'-TGGTCA-3') or with a single mismatched (MM) target sequence (5'-TGTTC-3') (single mutation underlined) (Fig. 2A and Fig. S6). As shown in Fig. 2, a single mutation in the binding sequence (MM) abolished Py-Im polyamide-induced pol II transcriptional arrest in the TS binding mode (both in the presence and absence of TFIIS). A similar effect is observed for the scaffold containing a single mutation in the NTS binding orientation (Fig. S6). Furthermore, DNA duplex melting experiments confirm that a single mutation significantly diminishes ligand binding affinity (Table S1), consistent with transcriptional inhibition results. These results demonstrate that polyamide-induced pol II transcriptional arrest is dependent upon the matched DNA sequence in both directions.

**Structural Insight into Py-Im Polyamide-Induced Pol II Arrest Revealed by Molecular Modeling.** To gain further mechanistic understanding of how pol II stalls upstream of the polyamide binding sites (2–5 bp), we modeled a series of pol II-arrested complexes in the presence of Py-Im polyamide binding to mimic pol II progression into the pausing region in a stepwise manner. The models are built based on the superposition of crystal structures of the downstream DNA duplex of pol II elongation complex (PDB ID code 3M3Y) (23) with the structure of the polyamide bound DNA duplex (PDB ID code 3I5L) (31). Here we modeled each of the four major pausing states observed in our biochemical studies. Details for model set-up can be found in *SI Materials and Methods*.





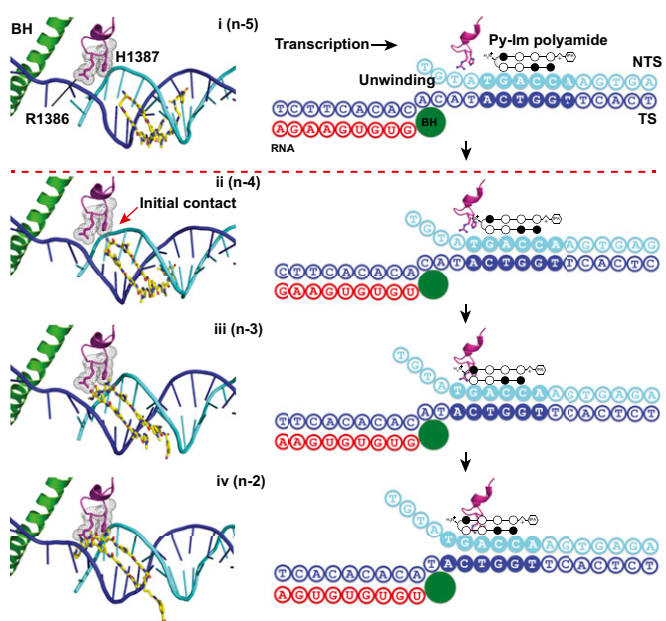
**Fig. 2.** Py-Im polyamides cause sequence-specific inhibition of pol II transcription elongation. The Py-Im polyamide binding site is in a TS orientation. (A) Scheme of single mutation of the designed target sequence. The PM sequence is 5'-WGGWCW-3', and the single-base MM sequence is 5'-WGWWCW-3' (mutation underlined). (B) Comparison of inhibition of **1** on pol II transcription elongation from the PM or MM sequence. The effect of TFIIIS was also investigated as shown in PAGE gel. Time points were 10 min, 20 min, 30 min, 1 h, and 2 h. Py-Im polyamide concentration was 0.5  $\mu$ M. (C) Quantitative analysis of selective transcriptional inhibition by **1**.

These models provide structural insights into the molecular mechanism of polyamide-induced pol II transcriptional arrest. First, as shown in Fig. S7A, the polyamide binding site remains within the intact duplex DNA region several base pairs downstream of the transcription bubble edge when pol II pauses at the first couple of pausing sites (states *i* and *ii*, or *n*-5 and *n*-4). This result suggests that pol II senses polyamide binding and pauses several base pairs before the actual polyamide binding site reaches the transcription bubble downstream edge. This result may rule out the simple model that pol II blockage results exclusively from prevention of the unwinding of DNA duplex at the transcription bubble downstream edge by polyamides. From high-resolution X-ray studies, it is known that the polyamide widens the minor groove of DNA, increasing the distance between the two DNA phosphodiester backbones (Fig. S7B) (31). These structural changes may cause misalignment of pol II residues involved in interactions with the DNA minor groove or phosphodiester backbone.

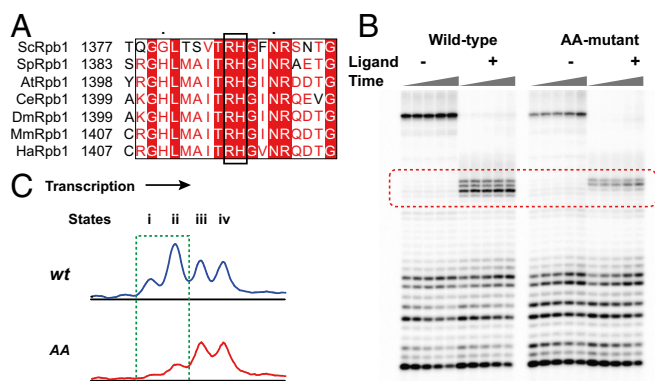
Structural analysis further reveals that the Switch 1 region in the Rpb1 subunit (38, 39) is likely to be involved in the interactions with polyamide-bound DNA. The Switch 1 region is a highly conserved motif of pol II that is located at the base of the clamp and serves as a hinge for clamp closure in the presence of a DNA–RNA hybrid in the active center (39). We identified two critical residues in the Switch 1 region of Rpb1, Arg1386 and His1387, which insert into the DNA minor groove in the wild-type pol II elongation complex and are likely either directly interacting with polyamide backbones (state *ii*, or *n*-4, for example) or sterically clashing with the Py-Im polyamide (states *iii* and *iv*) during RNA pol II translocation toward the Py-Im polyamide binding site (Fig. 3). Therefore, these two residues may play an interrogative role in sensing or scanning the DNA minor groove conformation and are likely to contribute to polyamide-induced pol II stalling.

Indeed, these models are consistent with the data that revealed two to four pausing bands with increasing difficulty to reach beyond the *n*-2 position (severe steric clash in state *iv*). Similar steric clash is expected between these two residues and the polyamide backbone for all four compounds regardless of their binding orientations (TS or NTS).

**Conserved Pol II Switch 1 Motif Functions as an Inspector for the DNA Minor Groove.** Our structural modeling revealed that two critical residues in the Switch 1 region of Rpb1 (Arg1386 and His1387) may serve as a sensor for the DNA minor groove and could stall pol II elongation via steric clash with minor groove binders. These two residues are conserved across eukaryotes from yeast to human (Fig. 4A). To verify the role of these two residues, we generated pol II with site-specific substitutions of Rpb1, R1386A and H1387A (AA-mutant) (see *SI Materials and Methods* for details) and performed transcription elongation assays in the presence and absence of hairpin **1** for both wild-type pol II and the AA-mutant. We theorized that the AA-mutant would have ablated steric clash with minor groove binders in early pausing states *i* and *ii*. Indeed, as shown in the dashed box in Fig. 4B, the pausing pattern of the AA-mutant is significantly different from that of the wild-type. The two stalled bands corresponding to states *i* and *ii* (Fig. 4B) were absent in the AA-mutant, revealing that these two residues are responsible for the early inspection of the downstream DNA minor groove environment before DNA unwinding. The downstream pol II pausing (states *iii* and *iv* or *n*-2 and *n*-3 position) was still observed upon polyamide binding, which was most likely caused by a direct blockage to DNA unwinding. Quantitative analysis (Fig. 4C) clearly indicated distinct blockage behaviors of the AA-mutant, in which the pausing at states *i* and *ii* (band *n*-4 and *n*-5) was abolished. This pattern



**Fig. 3.** Molecular modeling of RNA pol II transcriptional pausing and arrest by Py-Im polyamide. The functional interplay between two critical residues (R1386 and H1387) from the conserved Switch 1 region of Rpb1 and the polyamide-bound minor groove during RNA pol II progression. During translocation, the pol II may move from step *i* to step *iv* with distinct steric interactions. In step *i*, the Py-Im polyamide and these two residues are in positions without obvious steric interaction; in step *ii*, the polyamide and the residues establish direct contact; in steps *iii* and *iv*, strong steric clashes were observed between pol II Switch 1 residues and Py-Im polyamide, and duplex unwinding was also prevented, both of which lead to strong inhibition for RNA pol II elongation complex.



**Fig. 4.** Effects of Py-Im polyamide on the elongation of wild-type and AA-mutant are distinct. (A) Sequence alignment of Switch 1 region of Rpb1 of RNA pol II in different species. (B) Gel data of transcriptional elongation of the wild-type and AA-mutant. Time points were 10 min, 20 min, 30 min, 1 h, and 2 h. The different pausing patterns in the presence of 0.5  $\mu$ M Py-Im polyamide 1 are highlighted in the red dashed box. (C) Quantitative analysis of pausing patterns between the wild-type and AA-mutant. Different intensity peaks refer to different pausing states as described in Fig. 3. The major difference was highlighted in the green dashed box.

difference between the wild-type and AA-mutant confirmed the inspective role of Arg1386 and His1387 of the Rpb1 switch 1 motif in sensing the DNA minor groove before duplex unwinding. In addition, it is worth mentioning that compared with the wild-type pol II, the AA-mutant has a relatively strong sequence pausing effect during elongation (see the bottom part of gel in Fig. 4B), which may indicate possible roles of these two residues during normal transcriptional progression.

## Discussion

### Minor-Groove Sensor Motif Signals Transcription Pause and Blockage.

Although extensive mechanistic studies have been focused on how pol II recognizes and processes covalent DNA lesions (such as UV damage), or bypasses nucleosome barriers, less is known about the molecular mechanism of how pol II processes noncovalent DNA-binding small molecules. In this study, we investigated the effect of site-specific binding by Py-Im polyamides on pol II elongation in two binding orientations using biochemical assays, modeling, and mutation studies.

Our data from these studies revealed that an elongating pol II can sense polyamide binding, and pause as a stable complex immediately upstream of the binding site for over 20 h. This event is sequence-specific, where a single base pair mutation in the polyamide target sequence leads to the abolition of pol II arrest (Fig. 2). We further show the persistent pol II arrest induced by Py-Im polyamides cannot be rescued by TFIIS. This finding is remarkable given the fact that pol II itself has a strong capability to unwind the downstream DNA duplex to form the transcription bubble. This result is also in sharp contrast with the fact that TFIIS is able to rescue nucleosome-induced pol II arrest and a variety of pol II pausing/arrest events (13, 16, 40), considering that the binding affinity of histone–DNA interaction is very strong, within the similar nanomolar range as polyamides (41, 42). This observation may be explained by the kinetics of Py-Im polyamide DNA binding; the rate of dissociation is slow, comparable to DNA binding proteins, but the rate of association is much faster (43). Hence, the resistance to the effect of TFIIS may be attributed to the long-lifetime bound state of Py-Im polyamides during transcriptional elongation. These results support the idea that polyamides can block pol II elongation in a sequence-specific manner.

Furthermore, we find that substitutions of the functional moieties of Py-Im polyamides cause varied inhibitory potencies

toward pol II elongation with  $K_i$  values in the 1- to 100-nM range (Table 1). Intriguingly, the magnitude of inhibitory effects does not correlate directly with the polyamides' DNA binding affinities. This difference may reflect the difference between the aqueous solution where DNA binding affinity is measured and the amino acid-rich environment of the pol II downstream main channel where transcription inhibition is measured (Fig. S7).

Additionally, our results revealed that pol II senses polyamide binding in the minor groove via its conserved Switch 1 region and becomes stalled 2–5 bp upstream. The early polyamide-induced blockage of pol II elongation (n-4, n-5) is because of a steric clash with minor groove residues Arg1386 and His1387 during transcriptional progression toward the binding site (before formation of the transcriptional bubble), whereas the latter arrest (n-2, n-3) is attributed by direct prevention of DNA unwinding (Figs. 3 and 4 and Fig. S7).

**Biological Implications and Consequences.** Unlike DNA damaging agents, such as cisplatin, that covalently modify DNA in a global manner, Py-Im polyamides bind to the DNA minor groove noncovalently in a sequence-specific manner. Previous studies have indicated that polyamides can inhibit pol II elongation and trigger the degradation of Rpb1 *in vivo* (35). Our studies provide a mechanistic explanation of these observations. We found that polyamide-induced pol II arrest is stable for hours and cannot be rescued by TFIIS. Such prolonged stalled pol II is likely a target for ubiquitination and degradation in a similar manner to pol II stalled by cisplatin-induced DNA damage, and subsequent apoptosis (44, 45). The molecular processing by pol II transcription machinery may provide important insights into the understanding of how cells respond to this kind of noncovalent DNA minor groove binder *in vivo*. Given the potent sequence-specific inhibition by polyamide interference, and the demonstrated antitumor activity in xenografts with low host toxicity, it would be very attractive to consider whether there is a role in cancer therapy for these molecules. Transcription inhibition has been explored as a therapeutic strategy in cancer treatment (46–49). Inhibition of a ubiquitous and essential pathway in cancer could overcome, in part, the drug resistance conferred through the genetic heterogeneity of the disease. Indeed, a recent study has shown that many cancers have a concomitant hemizygous loss of *POLR2A* (*RPB1*) and p53, rendering them more vulnerable to transcription inhibition (49). Development of small molecules that selectively obstruct the pol II transcription machinery in a sequence-specific manner would present a therapeutic strategy to target transcription addition in cancer cells.

Finally, this study reveals a sensor function of the Switch 1 region to detect downstream DNA minor groove variations during pol II elongation. A highly conserved motif in pol II (Arg1386 and His1387 of Rpb1) has an inspective role within the DNA minor groove and can sense DNA minor groove binders. During transcription elongation, this motif likely interrogates the downstream duplex DNA before the formation of the transcription bubble. This motif and our previously identified major groove “epi-loop” (50) both exhibit DNA-sensing roles. Working together, these motifs can detect abnormalities in the downstream DNA template from both the major and minor grooves, which may play important roles in terms of controlling transcriptional fidelity and DNA damage recognition. Future studies may aim to reveal the function roles of this motif in the normal pol II transcription process, which may possibly unveil new insights into RNA pol II transcription regulation.

## Materials and Methods

**RNA Pol II Purification and *In Vitro* Transcription Assay.** Wild-type RNA pol II was purified from *Saccharomyces cerevisiae* as previously described, unless otherwise stated (51, 52). For the transcription inhibition assay, the RNA/DNA scaffold is preformed and incubated with various Py-Im polyamides for 3 h before the addition of pol II.



**DNA Template Design.** DNA sequences of template strand used for TS and NTS binding orientation were 5'-GACTCTTCTGACTTGGTCATACACACTTCTCTCTCGTTGTTCTCTCGATTGTAAGTGATGTGCTTGTAAAT-3' and 5'-GACTCTTCTGACTTGGTCATACACACTTCTCTCTCGTTGTTCTCTCGATTGTAAGTGATGTGCTTGTAAAT-3' (1–4 binding site is underlined).

**Py-Im Polyamide Synthesis.** Compounds 1–4 were synthesized on solid support as previously described (53–55). All compounds were purified by RP-HPLC and correct masses were verified by MALDI-TOF.

Additional experimental details including RNA pol II assembly, DNA template design, Py-Im polyamide synthesis, and molecular modeling can be found in *SI Materials and Methods*.

**ACKNOWLEDGMENTS.** This work was supported in part by NIH Grants GM102362 (to D.W.) and GM27681 (to P.B.D.); D.G. and J.N.S. were supported in part by the Intramural Research Program of the NIH, National Cancer Institute, and the Center for Cancer Research; and A.A.H. was supported in part by a postdoctoral Fellowship F32CA173977 from the NIH.

- Cheung AC, Cramer P (2012) A movie of RNA polymerase II transcription. *Cell* 149(7):1431–1437.
- Svetlov V, Nudler E (2013) Basic mechanism of transcription by RNA polymerase II. *Biochim Biophys Acta* 1829(1):20–28.
- Liu X, Bushnell DA, Kornberg RD (2013) RNA polymerase II transcription: structure and mechanism. *Biochim Biophys Acta* 1829(1):2–8.
- Selth LA, Sigurdsson S, Sveistrup JQ (2010) Transcript elongation by RNA polymerase II. *Annu Rev Biochem* 79:271–293.
- Dedrick RL, Kane CM, Chamberlin MJ (1987) Purified RNA polymerase II recognizes specific termination sites during transcription in vitro. *J Biol Chem* 262(19):9098–9108.
- Kerppola TK, Kane CM (1988) Intrinsic sites of transcription termination and pausing in the c-myc gene. *Mol Cell Biol* 8(10):4389–4394.
- Kellinger MW, et al. (2012) 5-formylcytosine and 5-carboxylcytosine reduce the rate and substrate specificity of RNA polymerase II transcription. *Nat Struct Mol Biol* 19(8):831–833.
- Xu L, et al. (2015) Impact of template backbone heterogeneity on RNA polymerase II transcription. *Nucleic Acids Res* 43(4):2232–2241.
- Tornaletti S, Patrick SM, Turchi JJ, Hanawalt PC (2003) Behavior of T7 RNA polymerase and mammalian RNA polymerase II at site-specific cisplatin adducts in the template DNA. *J Biol Chem* 278(37):35791–35797.
- Brueckner F, Hennecke U, Carell T, Cramer P (2007) CPD damage recognition by transcribing RNA polymerase II. *Science* 315(5813):859–862.
- Hardenbol P, Van Dyke MW (1992) In vitro inhibition of c-myc transcription by mitramycin. *Biochem Biophys Res Commun* 185(2):553–558.
- Mote J, Jr, Ghanouni P, Reines D (1994) A DNA minor groove-binding ligand both potentiates and arrests transcription by RNA polymerase II. Elongation factor SII enables readthrough at arrest sites. *J Mol Biol* 236(3):725–737.
- Kireeva ML, et al. (2005) Nature of the nucleosomal barrier to RNA polymerase II. *Mol Cell* 18(1):97–108.
- Saeki H, Sveistrup JQ (2009) Stability, flexibility, and dynamic interactions of colliding RNA polymerase II elongation complexes. *Mol Cell* 35(2):191–205.
- Hobson DJ, Wei W, Steinmetz LM, Sveistrup JQ (2012) RNA polymerase II collision interrupts convergent transcription. *Mol Cell* 48(3):365–374.
- Charlet-Berguerand N, et al. (2006) RNA polymerase II bypass of oxidative DNA damage is regulated by transcription elongation factors. *EMBO J* 25(23):5481–5491.
- Lagerwerf S, Vrouwe MG, Overmeer RM, Foustieri MI, Mullenders LH (2011) DNA damage response and transcription. *DNA Repair (Amst)* 10(7):743–750.
- Hanawalt PC, Spivak G (2008) Transcription-coupled DNA repair: Two decades of progress and surprises. *Nat Rev Mol Cell Biol* 9(12):958–970.
- Walmacq C, et al. (2015) Mechanism of RNA polymerase II bypass of oxidative cyclopurine DNA lesions. *Proc Natl Acad Sci USA* 112(5):E410–E419.
- Damsma GE, Cramer P (2009) Molecular basis of transcriptional mutagenesis at 8-oxoguanine. *J Biol Chem* 284(46):31658–31663.
- Cline SD, Riggins JN, Tornaletti S, Marnett LJ, Hanawalt PC (2004) Malondialdehyde adducts in DNA arrest transcription by T7 RNA polymerase and mammalian RNA polymerase II. *Proc Natl Acad Sci USA* 101(19):7275–7280.
- Jung Y, Lippard SJ (2006) RNA polymerase II blockage by cisplatin-damaged DNA. Stability and polyubiquitylation of stalled polymerase. *J Biol Chem* 281(3):1361–1370.
- Wang D, Zhu G, Huang X, Lippard SJ (2010) X-ray structure and mechanism of RNA polymerase II stalled at an antineoplastic monofunctional platinum-DNA adduct. *Proc Natl Acad Sci USA* 107(21):9584–9589.
- Kellinger MW, Park GY, Chong J, Lippard SJ, Wang D (2013) Effect of a monofunctional phenanthriplatin-DNA adduct on RNA polymerase II transcriptional fidelity and translation synthesis. *J Am Chem Soc* 135(35):13054–13061.
- Gottesfeld JM, Neely L, Trauger JW, Baird EE, Dervan PB (1997) Regulation of gene expression by small molecules. *Nature* 387(6629):202–205.
- Dickinson LA, et al. (1998) Inhibition of RNA polymerase II transcription in human cells by synthetic DNA-binding ligands. *Proc Natl Acad Sci USA* 95(22):12890–12895.
- Majumder P, et al. (2013) Minor groove binder distamycin remodels chromatin but inhibits transcription. *PLoS One* 8(2):e57693.
- Trauger JW, Baird EE, Dervan PB (1996) Recognition of DNA by designed ligands at subnanomolar concentrations. *Nature* 382(6591):559–561.
- White S, Szewczyk JW, Turner JM, Baird EE, Dervan PB (1998) Recognition of the four Watson-Crick base pairs in the DNA minor groove by synthetic ligands. *Nature* 391(6666):468–471.
- Kielkopf CL, et al. (1998) A structural basis for recognition of A.T and T.A base pairs in the minor groove of B-DNA. *Science* 282(5386):111–115.
- Chenoweth DM, Dervan PB (2009) Allosteric modulation of DNA by small molecules. *Proc Natl Acad Sci USA* 106(32):13175–13179.
- Hsu CF, et al. (2007) Completion of a programmable DNA-binding small molecule library. *Tetrahedron* 63(27):6146–6151.
- Nickols NG, Dervan PB (2007) Suppression of androgen receptor-mediated gene expression by a sequence-specific DNA-binding polyamide. *Proc Natl Acad Sci USA* 104(25):10418–10423.
- Muzikar KA, Nickols NG, Dervan PB (2009) Repression of DNA-binding dependent glucocorticoid receptor-mediated gene expression. *Proc Natl Acad Sci USA* 106(39):16598–16603.
- Yang F, et al. (2013) Antitumor activity of a pyrrole-imidazole polyamide. *Proc Natl Acad Sci USA* 110(5):1863–1868.
- Edelson BS, et al. (2004) Influence of structural variation on nuclear localization of DNA-binding polyamide-fluorophore conjugates. *Nucleic Acids Res* 32(9):2802–2818.
- Belotserkovskii BP, Mirkin SM, Hanawalt PC (2013) DNA sequences that interfere with transcription: Implications for genome function and stability. *Chem Rev* 113(11):8620–8637.
- Cramer P, Bushnell DA, Kornberg RD (2001) Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science* 292(5523):1863–1876.
- Gnatt AL, Cramer P, Fu J, Bushnell DA, Kornberg RD (2001) Structural basis of transcription: An RNA polymerase II elongation complex at 3.3 Å resolution. *Science* 292(5523):1876–1882.
- Kuraoka I, et al. (2007) RNA polymerase II bypasses 8-oxoguanine in the presence of transcription elongation factor TFIIS. *DNA Repair (Amst)* 6(6):841–851.
- Cotton RW, Hamkalo BA (1981) Nucleosome dissociation at physiological ionic strengths. *Nucleic Acids Res* 9(2):445–457.
- Ausio J, Seger D, Eisenberg H (1984) Nucleosome core particle stability and conformational change. Effect of temperature, particle and NaCl concentrations, and crosslinking of histone H3 sulfhydryl groups. *J Mol Biol* 176(1):77–104.
- Baliga R, et al. (2001) Kinetic consequences of covalent linkage of DNA binding polyamides. *Biochemistry* 40(1):3–8.
- Somesh BP, et al. (2005) Multiple mechanisms confining RNA polymerase II ubiquitylation to polymerases undergoing transcriptional arrest. *Cell* 121(6):913–923.
- Wilson MD, Harreman M, Sveistrup JQ (2013) Ubiquitylation and degradation of elongating RNA polymerase II: The last resort. *Biochim Biophys Acta* 1829(1):151–157.
- Stellrecht CM, Chen LS (2011) Transcription inhibition as a therapeutic target for cancer. *Cancers (Basel)* 3(4):4170–4190.
- Kwiatkowski N, et al. (2014) Targeting transcription regulation in cancer with a covalent CDK7 inhibitor. *Nature* 511(7511):616–620.
- Wang Y, et al. (2015) CDK7-dependent transcriptional addiction in triple-negative breast cancer. *Cell* 163(1):174–186.
- Liu Y, et al. (2015) TP53 loss creates therapeutic vulnerability in colorectal cancer. *Nature* 520(7549):697–701.
- Wang L, et al. (2015) Molecular basis for 5-carboxycytosine recognition by RNA polymerase II elongation complex. *Nature* 523(7562):621–625.
- Wang D, Bushnell DA, Westover KD, Kaplan CD, Kornberg RD (2006) Structural basis of transcription: Role of the trigger loop in substrate specificity and catalysis. *Cell* 127(5):941–954.
- Wang D, et al. (2009) Structural basis of transcription: Backtracked RNA polymerase II at 3.4 angstrom resolution. *Science* 324(5931):1203–1206.
- Dose C, Farkas ME, Chenoweth DM, Dervan PB (2008) Next generation hairpin polyamides with (R)-3,4-diaminobutyric acid turn unit. *J Am Chem Soc* 130(21):6859–6866.
- Puckett JW, Green JT, Dervan PB (2012) Microwave assisted synthesis of Py-Im polyamides. *Org Lett* 14(11):2774–2777.
- Li BC, Montgomery DC, Puckett JW, Dervan PB (2013) Synthesis of cyclic Py-Im polyamide libraries. *J Org Chem* 78(1):124–133.
- Irvin JD, et al. (2014) A genetic assay for transcription errors reveals multilayer control of RNA polymerase II fidelity. *PLoS Genet* 10(9):e1004532.
- Kireeva ML, Lubkowska L, Komissarova N, Kashlev M (2003) Assays and affinity purification of biotinylated and nonbiotinylated forms of double-tagged core RNA polymerase II from *Saccharomyces cerevisiae*. *Methods Enzymol* 370:138–155.
- Kellinger MW, Ulrich S, Chong J, Kool ET, Wang D (2012) Dissecting chemical interactions governing RNA polymerase II transcriptional fidelity. *J Am Chem Soc* 134(19):8231–8240.
- Xu L, Plouffe SW, Chong J, Wengel J, Wang D (2013) A chemical perspective on transcriptional fidelity: Dominant contributions of sugar integrity revealed by unlocked nucleic acids. *Angew Chem Int Ed Engl* 52(47):12341–12345.
- Xu L, et al. (2014) Strand-specific (asymmetric) contribution of phosphodiester linkages on RNA polymerase II transcriptional efficiency and fidelity. *Proc Natl Acad Sci USA* 111(32):E3269–E3276.
- Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2126–2132.
- Abagyan R, Totrov M, Kuznetsov D (1994) ICM—A new method for protein modeling and design: Applications to docking and structure prediction from the distorted native conformation. *J Comput Chem* 15:488–506.